

Developing Mobile Multimodal Applications

Raimund Schatz, Rainer Simon, Hermann Anegg, Florian Wegscheider, Georg Niklfeld
Telecommunications Research Center Vienna (ftw.), Donau-City-Straße 1, 1220 Wien, Austria
{schatz, simon, anegg, wegscheider, niklfeld}@ftw.at

In this paper we present three major aspects of our research on mobile multimodal applications and middleware: device-independence, prototype applications, and user involvement. We followed model-based principles, which provided us with challenges which we report upon in this paper. We discuss the development of a multimodal game and a messaging client designed for mobile hands-busy or eyes-busy situations. We involved users throughout the whole development process in order to gain insights into the appropriateness and possible use of different modalities in various contexts.

Keywords: mobile applications and devices, multimodal user interfaces, device independence, user centred design.

1. INTRODUCTION

Today there are still only a few applications for mobile terminals that exploit several modalities (e.g. voice and graphics) in a flexible manner, even though the generally small screen and frequent context-changes suggest that a multimodal interface yields great user benefits. MONA (Mobile multimodal Next-generation Applications, <http://mona.ftw.at>) focuses on middleware for multimodal applications targeted at a range of different mobile client devices such as PDAs and Smart-Phones. Users should be able to interact visually, by voice or a combination of both according to their individual preferences. To this end, we developed a presentation server which renders the application's abstract UI-descriptions to multimodal device-specific interfaces (Fig. 1).

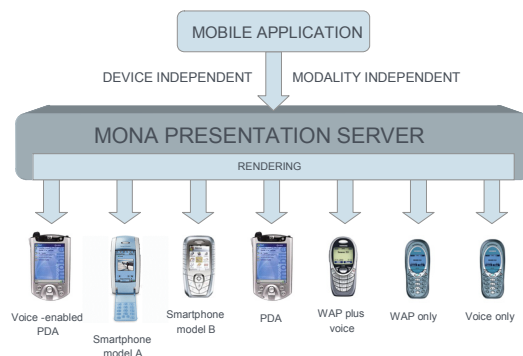


FIGURE 1: MONA System Architecture.

2. DEVICE-INDEPENDENCE AND AUTHORING

We defined a proprietary XML user interface description language, consisting of a platform-independent widget set and elements inspired by model-based approaches [3]. However, we have found that the abstract nature of our approach and the high degree of automation during the user interface generation process was difficult to combine with the work practices of our designers: rather than working in a top-down fashion, their workflow was highly iterative. Also, they made use of visual tools and artefacts (such as GUI sketches) in order to precisely predict the appearance of the UI on the different devices they were designing for. Model-based authoring approaches cannot provide the high degree of predictability that designers are accustomed to. This has also been acknowledged by others [2]. To enable predictability and allow for a smooth transition from current design practices, we have developed the MONAEditor – a prototype authoring tool featuring real-time GUI emulations for different devices and live audio preview. A real-time work environment such as this not only enhances productivity but also fosters experimentation with single-authoring, a domain that is still new for a large part of the interface design community.

3. DEMO-APPLICATIONS

We developed two prototype applications for two usage domains that we think can highly benefit from mobile multimodality: entertainment and work. Instead of adapting existing applications we designed both applications from scratch in order to let the voice and the visual aspects of the interface truly co-evolve.

Mona@play is a multimodal quiz game in the style of “Who wants to be a millionaire?”. Rich interaction between players significantly enhances the entertainment value of multi-user games. To this end, mona@play features multimodal chat which complements the competitive game play and supports asymmetric multimodality: spoken chat messages are automatically translated to text messages for GUI-only receivers with a conversational ASR (automatic speech recognition) subsystem and vice versa. In addition, each user is represented by an avatar with context-dependent moods, one of the requirements that surfaced in initial participatory design workshops (Fig. 2).



FIGURE 2: The quiz game application as rendered on PDA, Smart- and WAP-phone and the messaging client (right).

Mona@work is a mobile multimodal messaging client. It enables the user to manage different types of text and voice messages through an interface designed to support mobile users who frequently switch between visual and voice modalities (e.g. while on the go). Important functions can be quickly accessed by a voice-only interface (e.g. the reading out of new messages via TTS).

4. DEVELOPMENT PROCESS AND USER ASPECTS

We undertook user studies at each stage of the project utilising various techniques [1] borrowed from Scenario and Participatory Design. User involvement was essential for finding out more about the potential use of multimodality within mobile applications and their actual usage in context: evaluating in the laboratory and in the field, we found that in the lab context most users did not find voice interaction useful. However, this result was completely reversed when the users were asked the same question after using the application in the field (with the system being rated as 3.84 in the field and 2.44 in the lab (1= poor, 5= excellent)). We also found that in the lab the users preferred to use either voice or text and did not mix modalities, whereas in the field the users tended to mix the modalities.

5. CONCLUSIONS AND FUTURE WORK

Our research shows that multimodality can truly enhance mobile applications, particularly in mobile contexts: our test users appreciated the freedom to use alternative modalities. We also learned that application designers should not treat the voice modality as an add-on to the visual interface. In order to support true multimodality, the voice interaction must co-evolve with the visual interface design. Therefore, we believe that more research on tools and methods for speech-recognition, authoring and for carrying out field studies for multimodal mobile application is crucial if we are to gain user acceptance of multimodality.

6. ACKNOWLEDGEMENTS

This work is part of the projects MONA and SUPRA funded by the Austrian competence centre programme Kplus and the companies Kapsch CarrierCom, Mobilkom Austria, and Siemens Austria. Kirusa Inc. provided their multimodal platform, SVOX Ltd its text-to-speech technology, and Nuance Inc its speech recognition technology.

REFERENCES

- [1] Baillie, L., Benyon, D.R., Macaulay, C., Petersen, M.G. (2003) Investigating Design Issues in Household Environments. *Cognition, Technology and Work*, Volume 1., Springer, London.
- [2] Myers, B., Hudson, S. E., Pausch, R. (2000) Past, Present, and Future of User Interface Software Tools. *ACM Transactions on Computer-Human Interaction 2000*, Volume 7, Issue 1, ACM, New York.
- [3] Florins, M., Vanderdonckt, J., Graceful Degradation of User Interfaces as a Design Method for Multiplatform Systems. *Proceedings of IUI 2004*, ACM, New York.